METHOD AND APPARATUS FOR DETERMINING THE IDENTITY OF A USER BY NARROWING DOWN FROM USER GROUPS

Field of the Invention

The present invention generally relates to user authentication and identification

5   methods, i.e. methods and apparatus for determining the identity of a user. The present

invention specifically relates to systems that recognize the identity of a user given a

biometric sample such as voice, fingerprint, hand geometry, iris, etc.

Background of the Invention

Current solutions to problems of the type just describe use one or more of the

10   following authentication/identification methods: possessing an id-device (e.g. door key),

knowing a certain piece of knowledge (e.g. passwords), and biometrics (e.g. voice print).

Biometrics have the advantageous property of using an inherent attribute of the user (e.g.

a fingerprint). Biometric systems perform user authentication and/or identification. For

example, a speaker verification system determines the identity of a person given their

15   speech sample. Unlike some other types of biometrics such as fingerprint recognition

(referred to as *static biometrics* herein), the more a person speaks, the better the voice can

be characterized and hence the higher the accuracy of the speaker recognition system;

biometrics that have this property are referred to herein as *dynamic* biometrics. Some

examples of static biometrics are: fingerprint, iris, retina, and hand geometry, while examples of dynamic biometrics include voice, gait, and keyboard stroke.

Dynamic biometrics systems such as speaker recognition systems exhibit reduced accuracy when less biometric data is available (for example when the user does not speak much). Therefore, such systems will typically try to elicit more data from the user, which is impractical in some applications. Whenever there is not enough data to make an accurate identity decision, current dynamic biometrics systems may simply fail to determine who the user is, without providing additional information that may characterize the user even without knowing her/his identity.

A need therefore has been recognized in connection with providing dynamic biometrics systems that improve upon the shortcomings of the efforts made to date.

Summary of the Invention

There is broadly contemplated, in accordance with at least one preferred embodiment of the present invention, the performance an authentication/identification task by narrowing down the possible class of user identities, in a refined fashion, as the user speaks, walks, types or performs some other function. For example, for a certain speaker recognition system 20 seconds of speech data might be required to accurately determine who the speaker is. However, it is recognized herein that, e.g., after 2 seconds

it is distinctly possible to accurately determine that the user is a female, and after an additional 5 seconds determine that it's a female in her 30's, after 6 more seconds determine that she has a southern accent, etc. In this way the system gradually *narrows down* the user's identity subset. Such an approach can represent part of a holistic user profiling system that is able to provide information about the user in an *incrementally* refined manner. It also permits a user to be recognized to some degree without the requirement of explicitly enrolling a model or template from the user's reference biometrics. Hence, low security transaction and related applications could be enabled through basic user profiling checks on the user.

In at least one preferred embodiment of the present invention, two components are used in concert:

1. A method/apparatus to characterize a user by his/her level of match with predetermined user groups (male/female, accent, age, fast walkers, slow walkers, voice quality, voice thickness, roughness, softness, speaking style). This is referred to herein as a *user profiler*.

2. A method/apparatus to compute a confidence measure reflecting how confident the system is that the user belongs to a particular user group (e.g. a measure representing

how confident the system is that a speaker is a male speaker). This is referred to herein as a *confidence estimator.*

The user profiler and the confidence estimator preferably use user-group models to determine their output vectors. For example, the user profiler may use user-group models trained on subsets of the user population such as: male, female, hoarse-voice, slow walkers, etc. Both the profiler and confidence estimator preferably operate as biometric data is being collected (i.e. as the user speaks/walks/types), and allow the user to be authenticated/verified in a "narrow down" process. In this process, the system gradually determines confidently that the user belongs to additional groups, until it potentially determines confidently who the user is. The process can be likened to an application of successive sieves that filter speaker characteristics with increasing precision.

In summary, one aspect of the present invention provides a method for assessing the identity of an individual, said method comprising the steps of: accepting input from an individual; attributing at least one user group to the individual; and repeating said attributing step until the identity of the individual is assessed.

An additional aspect of the present invention provides an apparatus for assessing the identity of an individual, said apparatus comprising: an arrangement for accepting

input from an individual; and an arrangement for attributing at least one user group to the individual; said attributing arrangement being adapted to repeat the attributing until the identity of the individual is assessed.

Furthermore, another aspect of the present invention provides a program storage
5   device readable by machine, tangibly embodying a program of instructions executable by the machine to perform method steps for assessing the identity of an individual, said method comprising the steps of: accepting input from an individual; attributing at least one user group to the individual; and repeating said attributing step until the identity of the individual is assessed.

10   For a better understanding of the present invention, together with other and further features and advantages thereof, reference is made to the following description, taken in conjunction with the accompanying drawings, and the scope of the invention will be pointed out in the appended claims.

Brief Description of the Drawings

15   Fig. 1 is a schematic block diagram of primary components in accordance with an embodiment of the present invention.

Fig. 2 is essentially the same diagram as Fig. 1 but illustrating an additional step.

Fig. 3 is a schematic block diagram depicting a second enrollment method.

Description of the Preferred Embodiments

Fig. 1 illustrates a system 100 configured in accordance with a preferred

5   embodiment of the present invention. A user's biometric sample 102 (such as speech) is

preferably input and fed to a user profiler 104 and confidence estimator 106, as described

further above. A user's group match scores 108 and group confidence scores 110,

respectively, are preferably provided as output.

Preferably, a speaker may enroll in the system in one of two ways. As one

10   possible measure, the user may provide biometric data (e.g. speak) while both the profiler

104 and confidence estimator 106 are operating. Once enough confidence measures are

met, there then will develop an indication that the user belongs to the corresponding user

groups. The match levels for the confident groups, represented by a vector of profiler

scores, then serve as the user's model/template that will be used as a reference when the

15   user's identity needs to be determined in the future. This is referred to as *enrollment*

method 1 herein. Fig. 2 schematically illustrates this method; it essentially is the same

illustration as Fig. 1 but shows an additional feed of confident user groups 112 between user confidence scores 110 and user group match scores 108.

As another possible measure, a profiler may be enhanced to include an additional group which includes only the user. When this method is used, user enrollment involves the same procedures that are used to enroll a user-group in the profiler and confidence estimator. This is referred to as *enrollment method 2* herein, and is illustrated in Fig. 3. Thus, with a biometric sample 202, a user group may be enrolled (214) at which point the resulting new user group 216 could be used in a user profiler 104 or confidence estimator 106 as in Fig. 1.

Generally, referring back to Fig. 1, when a user needs to be authenticated or identified, the user speaks/walks/types, the profiler 104 operates in an ongoing manner, and thus issues group-match scores 108. In parallel, the confidence estimator 106 issues group-confidence scores 110. Once a given confidence measure meets a threshold, the user is deemed to belong to the corresponding user group. The system then preferably issues a *cue*. The identity determination process (either authentication or identification) is thus preferably released as a series of cues over time. When sufficient data is available, the final cue may be the user's identity. The cues can make use of essentially information

conveyed in the biometric signal. For speech, this may be acoustic/spectral information, words, content, emotional cues, etc.

The embodiments of the present invention may be used for both user identification and authentication. For user identification, an example of returned cues during the time that a user speaks might be:

<male><between 25 and 45 years old><Has foreign accent><Breathy voice><nervous><likely to have college education><polite><speaks fast><John Smith>

For user authentication, with a target speaker class of "John Smith", an example of returned cues during the time the user speaks might be:

<Indeed a male><Age range found to match John's age><has breathy voice like John><It is John>

Or:

<female = NOT John>.

If the user enrolled using enrollment method 1, then authentication may be performed in the following way. Once the user provides enough biometric data such that

all of the groups she/he belongs to are confident (meet the confidence thresholds), a similarity score is computed as a distance measure between the vector of profiler match scores during authentication and during enrollment. This score is then thresholded to decide whether to accept the user's identity claim or reject it. Similarly, for user identification the system preferably computes profiler and confidence scores for all enrolled users. Once a confident profiler vector is obtained with respect to all enrolled users, and once the profiler vector of the test biometrics meets the confidence thresholds, the user's identity is determined to be the one corresponding to the user for which the distance measure between the test biometrics' profiler vector and the user vector is the smallest.

If the user enrolled using enrollment method 2, then authentication may be performed in the following way. Once the confidence score of the user model meets a threshold, a user authentication decision can be made by thresholding the score that the profiler produced for the user model. If the session ends prior to confident authentication of the user model, the partial confident information obtained for other models can be used.

Though enrollment methods 1 and 2 have been described hereinabove individually, it is certainly the case that a combination of both methods may also be used.

Though the manners and algorithms that could be employed for carrying out the embodiments of the present invention as described above are potentially vast, the algorithms described and contemplated in the following references have been found to be particularly meaningful in connection with different aspects of the present invention: for statistical modeling and Gaussian Mixture Models (GMM), G.N. Ramaswamy, J. Navratil, U.V. Chaudhari, R.D. Zilca, ``The IBM system for the NIST 2002 cellular speaker verification evaluation," ICASSP-2003, Hong Kong, April, 2003; and for discriminative methods such as Support Vector Machines (SVM), S. Fine, J. Navratil, R.A. Gopinath, ``A hybrid GMM/SVM approach to speaker Identification," ICASSP 2001, Salt Lake City, Utah, May 2001. The methods described in these two references are currently used to enroll user models in biometric systems, but can be used as-is to enroll user groups, simply by feeding the enrollment method with biometric data exclusively from a group of users instead of from a single user.

It is to be understood that the present invention, in accordance with at least one presently preferred embodiment, includes an arrangement for accepting input from an individual and an arrangement for attributing at least one user group to the individual. Together, these elements may be implemented on at least one general-purpose computer running suitable software programs. These may also be implemented on at least one

Integrated Circuit or part of at least one Integrated Circuit. Thus, it is to be understood that the invention may be implemented in hardware, software, or a combination of both.

If not otherwise stated herein, it is to be assumed that all patents, patent applications, patent publications and other publications (including web-based publications) mentioned and cited herein are hereby fully incorporated by reference herein as if set forth in their entirety herein.

Although illustrative embodiments of the present invention have been described herein with reference to the accompanying drawings, it is to be understood that the invention is not limited to those precise embodiments, and that various other changes and modifications may be affected therein by one skilled in the art without departing from the scope or spirit of the invention.